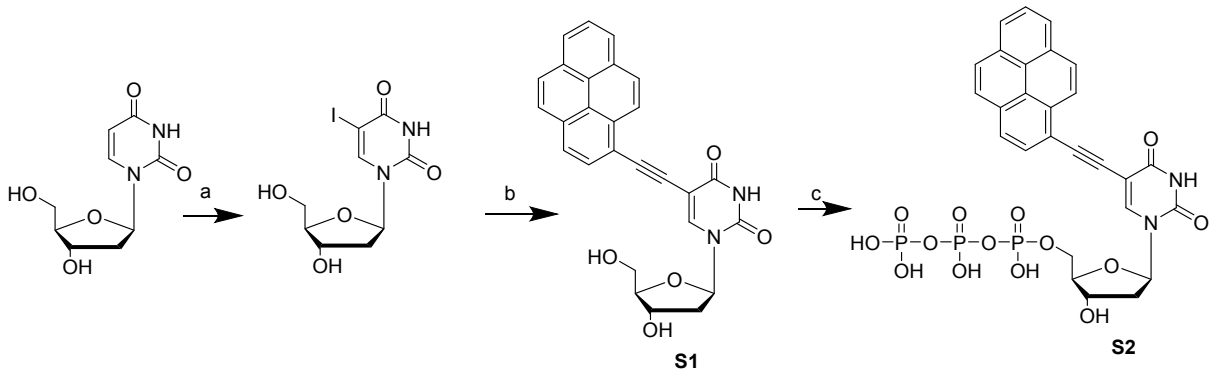
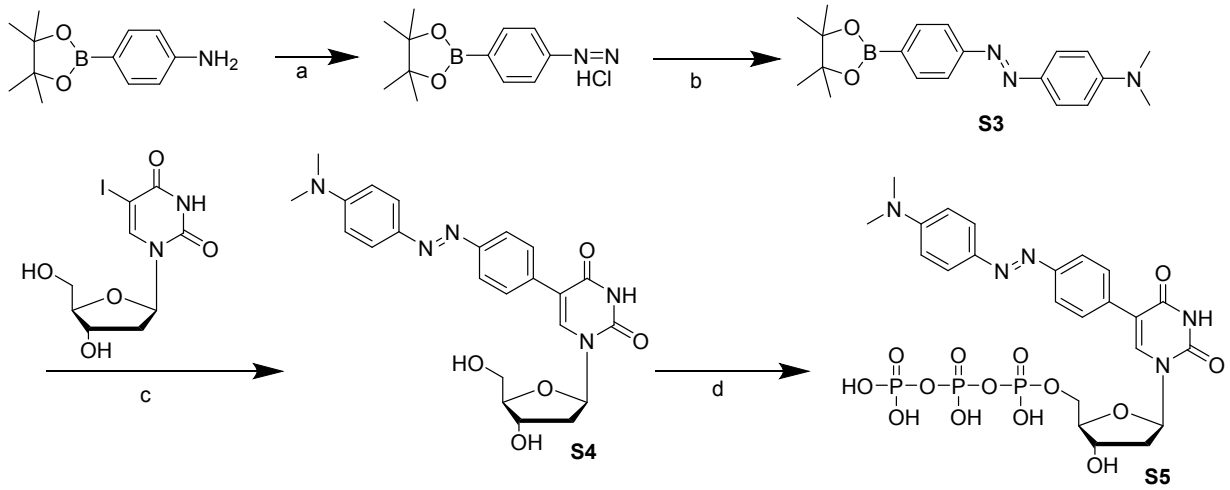
Supplementary Information

Site-specific incorporation of multiple units of functional nucleotides into DNA using a step-wise approach with polymerase and its application to monitoring DNA structural changes

Synthesis of fluorescent and quencher nucleotides



Scheme S1: Synthesis of the fluorescent nucleotide **dUpyrTP**. (a) Iodine, silver nitrate, MeOH, 40 °C, 3 h, 100%; (b) 1-ethynylpyrene, $(PPh_3)_4Pd$, CuI, DIPEA, DMF, 45–50 °C, 5 h, 75%; (c) proton sponge, $POCl_3$, Bu_3N, Bu_3NPPi, $(MeO)_3P$, $Bu_3NH_4P_2O_7$/DMF, $Et_3NH_2CO_3$, –20 °C, 3 h.



Scheme S2. Synthesis of the quencher nucleotide **dUazoTP**. (a) 4-Aminophenylboronic acid pinacol ester, $NaNO_2$, conc. HCl, 0–5 °C; (b) *N*,*N*-dimethylaniline, acetic acid, 0–5 °C; (c) $Pd(OAc)_2$, TPPTS, Na_2CO_3, H_2O/CH_3CN (1:2), 80%; (d) proton sponge, $POCl_3$, Bu_3N, Bu_3NPPi, $(MeO)_3P$, $Bu_3NH_4P_2O_7$/DMF, $Et_3NH_2CO_3$, –20 °C, 3 h, 50%.

General Procedure for Triphosphate Synthesis

Proton sponge (1.5 equiv) and free nucleoside (1 equiv) were dissolved in trimethylphosphate (0.3 M) and cooled to –20 °C. $POCl_3$ (1.5 equiv) was added dropwise, and then the purple slurry was stirred at –20 °C for 2 h. Tributylamine (6.2 equiv) was added, followed by a solution of tributylammonium

pyrophosphate (5.0 equiv) in DMF (0.5 M). After 5 min, the reaction was quenched by the addition of 0.5 M aqueous $Et_3NH_2CO_3$ (20 vol-equiv) and the resulting solution was lyophilized. Purification through reverse-phase (C18) HPLC (4–35% CH_3CN in 0.1 M $Et_3NH_2CO_3$, pH 7.5), followed by lyophilization, afforded the triphosphate as a solid.

Synthesis of the fluorescent nucleotide dUpyrTP

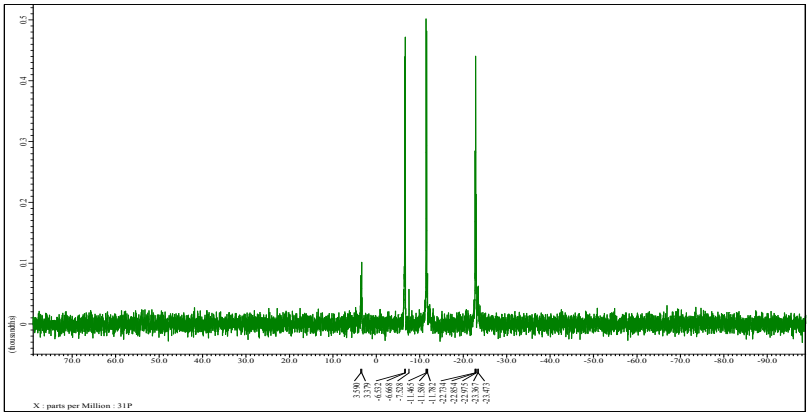
5-Ethynylpyrene-2′-deoxyuridine (S1)

CuI (21.5 mg, 0.110 mmol) and $(PPh_3)_4Pd$ (32.6 mg, 0.0280 mmol) were added to a solution of 5-iodo-2′-deoxyuridine (200 g, 0.564 mmol) in DMF (6 mL), and then the mixture was degassed. *N*,*N*-Diisopropylethylamine (0.196 mL, 1.10 mmol) and 1-ethynylpyrene (382 mg, 1.69 mmol) were added and then the yellow solution was stirred for 4 h at 55 °C under N_2. The reaction was monitored using TLC (CH_2Cl_2/MeOH, 10:1). Upon completion, the reaction mixture was diluted with CH_2Cl_2 (20 mL) and extracted with H_2O (2 × 15 mL). The organic phase was separated, dried (anhydrous Na_2SO_4), and evaporated to dryness. The residue was purified through column chromatography (SiO_2; CH_2Cl_2/MeOH, 10:1) using a Combiflash machine to afford a yellow solid (216 mg, 0.477 mmol, 84.6%). ^{1}H NMR (400 MHz, DMSO-d_6): δ = 13.14 (s, br, 1H; NH-1), 9.97 (s, 1H; H-6), 9.68–9.45 (m, 9H; pyrene), 7.51 (t, *J* = 10 Hz, 1H, H-1′), 6.61–6.64 (m, 2H; OH-3′, OH-5′), 5.64 (t, *J* = 10 Hz, 1H; H-3′), 5.10–4.90 (m, 2H; 2H-5′), 4.65 (s, 1H; H-4′), 3.59–3.42 (m, 2H; 2H-2′).

5-Ethynylpyrene-2′-deoxyuridine-5′-*O*-triphosphate (S2)

S1 (10 mg, 0.022 mmol) was transformed to the nucleotide-5′-*O*-triphosphate as described in the **General Procedure for Triphosphate Synthesis**. **^{31}P NMR** (400 MHz, D_2O): δ –22.734 to –23.473 (m; β-P), –11.465 to -11.782 (m,; α-P), –6.532 to -7.528 (m; γ-P). **^{1}H NMR** (400 MHz, D_2O): δ = 8.03 (s, 1H; H-6), 7.88–8.01 (m, 9H; pyrene), 6.74–6.70 (m, 1H; H-1′), 4.67–4.54 (m, 1H; H-3′), 4.33 (s, 1H; H-4′), 3.98–3.95 (m, 2H; H-5′), 280–2.70 (m, 2H; 2H-2′). **MS** (**LC/MS**, *m*/*z*): Calculated for $C_{27}H_{23}N_2O_{14}P_3$: 692.04; found: 691.1 ([M – H$^+$]).

^{31}P-NMR spectra of 5-Ethynylpyrene-2′-deoxyuridine-5′-*O*-triphosphate (S2)



Synthesis of the quencher nucleotide dUazoTP

***N*,*N*-Dimethyl-4-((4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)azo aniline (S3)**

A solution of 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)aniline (1.00 g, 4.56 mmol), water (10 mL), and concentrated HCl (1.5 mL) was cooled to 0–5 °C. A solution of sodium nitrite (330 mg, 4.78 mmol) in water (15 mL) was added dropwise, keeping the temperature below 10 °C. After 30 min, a solution of *N*,*N*-dimethylaniline (567 mg, 4.68 mmol) in acetic acid (1 mL) was added at 0–5 °C. The solution was stirred for 30 min at 0–5 °C and then warmed to room temperature. The product precipitated from the solution after neutralization (20% NaOH) to a pH of approximately 7. The

brown solid was filtered off and washed sequentially with cold water and cold MeOH. The product was purified through chromatography (SiO_2; CH_2Cl_2/MeOH, 95:5) to give a red powder (345 mg, 22%). **^{1}H NMR** ($CDCl_3$, 400 MHz, ppm): δ 7.95–7.83 (m, 4H), 7.82 (d, *J* = 7.4 Hz, 2H), 6.76 (d, *J* = 7.4 Hz, 2H), 3.09 (s, 6H), 1.36 (s, 12H).

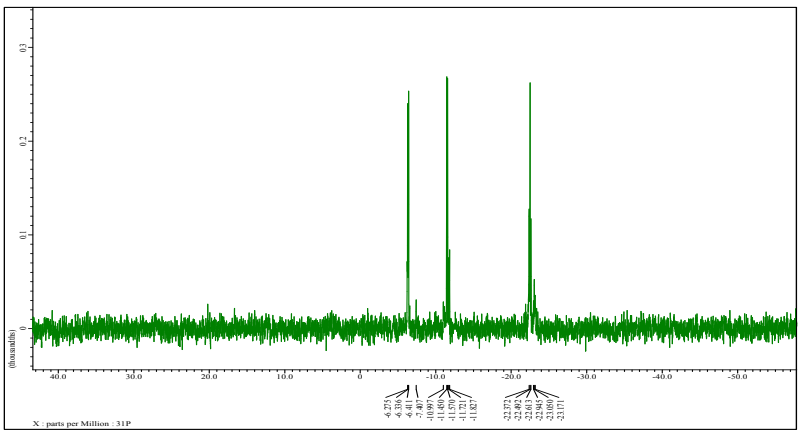
5-(8-(4-(4-(Dimethylamino)azobenzene)-2′-deoxyuridine (S4)

S3 (130 mg, 0.370 mmol), $Pd(OAc)_2$ (3.13 mg, 0.0140 mmol), TPPTS (39.8 mg, 0.0700 mmol), and Na_2CO_3 (90.0 mg, 0.846 mmol) were added to a solution of 5-iodo-2′-deoxyuridine (100 g, 0.282 mmol) in water/MeCN (2:1, 6 mL) under N_2. The mixture was stirred under reflux for 2 h and monitored using TLC (CH_2Cl_2/MeOH, 10:1). Upon completion, the reaction mixture was diluted with CH_2Cl_2 (20 mL) and extracted with H_2O (2 × 15 mL). The organic phase was separated, dried (anhydrous Na_2SO_4), and evaporated to dryness. The residue was purified through column chromatography (SiO_2; CH_2Cl_2/MeOH, 10:1) using a Combiflash machine to afford an orange solid (114 mg, 0.252 mmol, 89.4%). **^{1}H NMR (400** MHz, DMSO-d_6): δ 10.31 (s, br, 1H; NH), 8.39 (s, 1H; H-6), 7.80–7.75 (m, 6H; Ar-H), 6.83 (d, *J* = 15 Hz, 2H; Ar-H), 6.23 (t, *J* = 10 Hz, 1H; H-1'), 5.32–5.22 (m, 2H; OH-3′, OH-5′), 4.32 (s, br, 1H; H-3′), 3.84–3.83 (m, 1H; H-4′), 3.70–3.54 (m, 2H, CH_2-5′), 1.19 (t, *J* = 1.3 Hz, 6H; CH_3).

5-(8-(4-(4-(Dimethylamino)azobenzene)-2′-deoxyuridine-5′-*O*-triphosphate (S5)

S4 (10 mg, 0.022 mmol) was transformed to the nucleotide-5′-*O*-triphosphate as described in the **General Procedure for Triphosphate Synthesis**. **^{31}P NMR** (400 MHz, D_2O): δ –22.372 to –23.171 (m; β-P), –10.997 to -11.827 (m,; α-P), –6.275 to -6.411 (m; γ-P). **^{1}H NMR** (400 MHz, D_2O): δ 8.15 (s, 1H; H-6), 7.95–7.83 (m, 6H, ArH), 7.64–7.61 (m, 2H, ArH), 6.68 (m, 1H, H-1′), 4.50–4.40 (m, 1H, H-3′), 4.25–4.18 (m, 1H, H-4′), 3.80–3.72 (m, 2H, CH_2-5′), 2.08–2.04 (m, 2H, H-2′), 0.89 (t, 6H, CH_3). **MS (LC/MS**, *m*/*z*): Calculated for $C_{23}H_{28}N_5O_{14}P_3(C_2H_5)_3N^+$: 793.9, found: 792.21 ([M + Et_3N^+ – H^+]).

31P-NMR spectra of 5-(8-(4-(4-(Dimethylamino)azobenzene)-2′-deoxyuridine-5′-*O*-triphosphate (S5)



Photophysical Properties of Fluorescent Nucleotide

We used 9,10-diphenylanthracene in EtOH as a reference compound (ϕ_r = 0.95) to calculate the quantum yields of our fluorescent compounds.

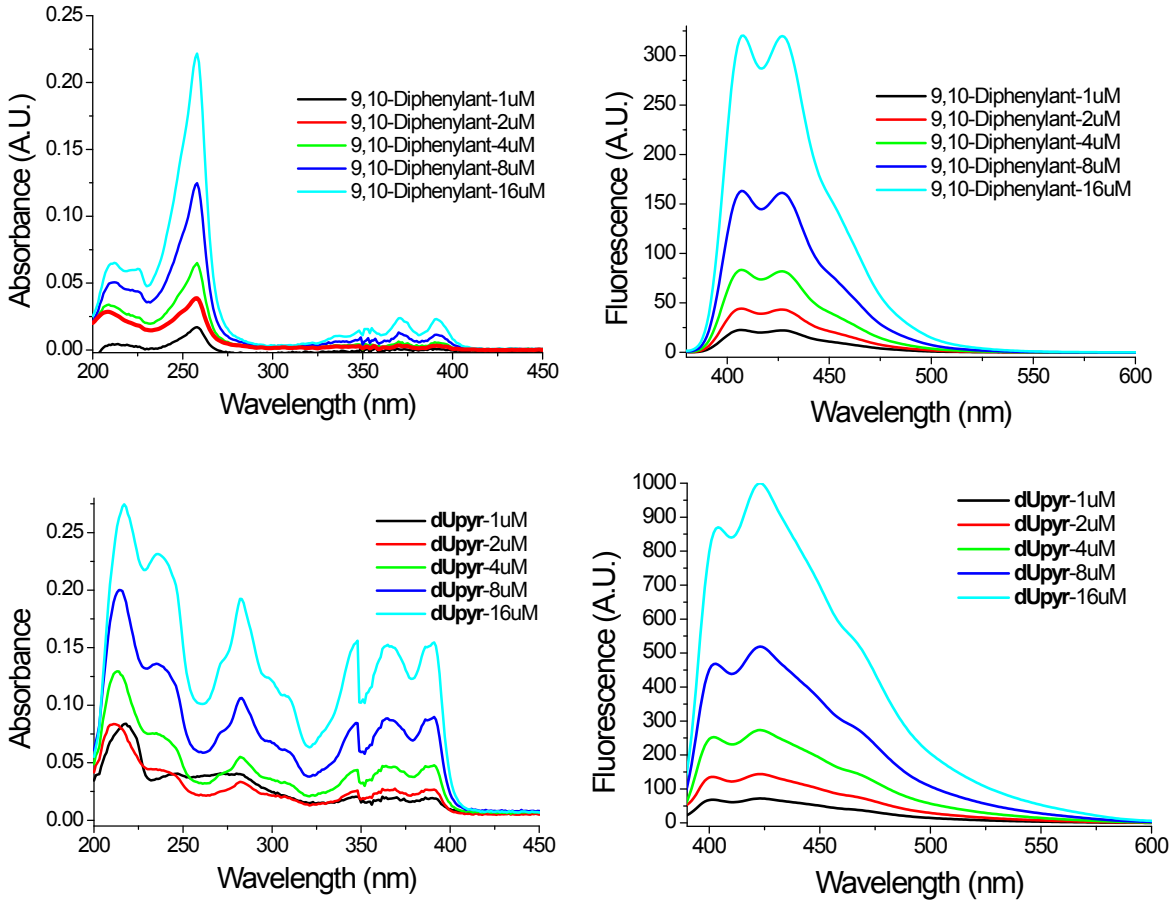


Figure S1. Absorption and fluorescence spectra of the fluorescent nucleotide and reference compounds, recorded at concentrations of 1–16 µM in EtOH at room temperature. Fluorescence emission spectra were recorded with excitation wavelengths for anthracene (374 nm) and pyrene (386 nm).

Quantum yields were also calculated.

Compounds / Impact	Reference	**dUpyrTP**
Gradient	14,927	13,450.01
Quantum yield	0.95	0.856

Table S1: Designed Oligonucleotides

Sequence	5′–3′
Primer **P1**	ATC CAA GGA TGC ATT
Template **T1**	[Phosphate] A CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG **A** AAT GCA TCC TTG GAT
Template **T2**	[Phosphate] A CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG **AA** AAT GCA TCC TTG GAT
Template **T3**	[Phosphate] A CTG CTG CTG CTG CTG CTG CTG CTG CTG CTG **AAA** AAT GCA TCC TTG GAT

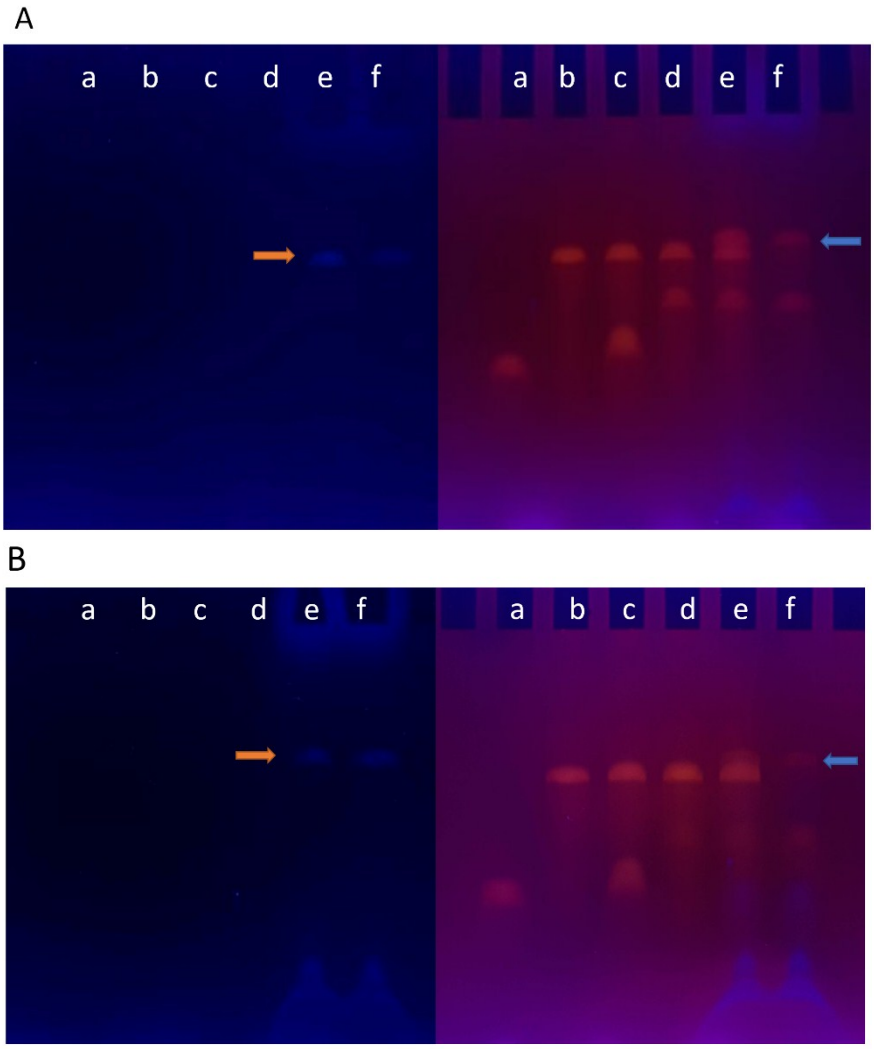


Figure S2. Denaturing PAGE characterization of the proposed strategy with the introduction of two **dUazoTP** units into the **P1:T2** duplex sequence (A) and three **dUazoTP** units (B) into the **P1:T3** duplex sequence. Fluorescence (left) and staining (EtBr solution; right) images the gel data. (a) **Primer P1**; (b) **Template T2** or **T3**; (c) after incorporating two or three **dUazoTP** units; (d) after extending with **dATP**, **dGTP**, and **dCTP**; e) full-length product after incorporation of **dUpyrTP**; (f) full-length product after treatmnt with Lambda exonuclease. **Primer P1** (0.1 mM) and **Template T2** or **T3** (0.1 mM) were annealed to form the **Primer P1:Template T2** (or **T3**) duplex by heating to 95 °C and then slowly cooling to room temperature. nPfu DNA polymerase (2 U) was added with **2 mM dUazoTP** and then the system was incubated at 37 °C for 12 h; this sample was purified using a QIAquick nucleotide removal kit (**step 1**). nPfu DNA polymerase (2 U) and 2 mM **dATP**, **dGTP**, and **dCTP** were added to this sample, which was then incubated at 37 °C for 12 h and purified using a QIAquick nucleotide removal kit (**step 2**). nPfu DNA polymerase (2 U) and 2 mM **dUpyrTP** were added to the products from **step 2** and then the system was incubated at 37 °C for 12 h and purified using a QIAquick nucleotide removal kit (**step 3**). The full-length product was treated with Lambda exonuclease (5U, Glycine/KOH buffer, 37 °C, 30 min) (**step 4**). All polymerase extensions were stopped by adding twice the amount of stop buffer (10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol, and 0.1% bromophenol blue in formamide). All product solutions were loaded onto a 20% denaturing polyacrylamide gel.

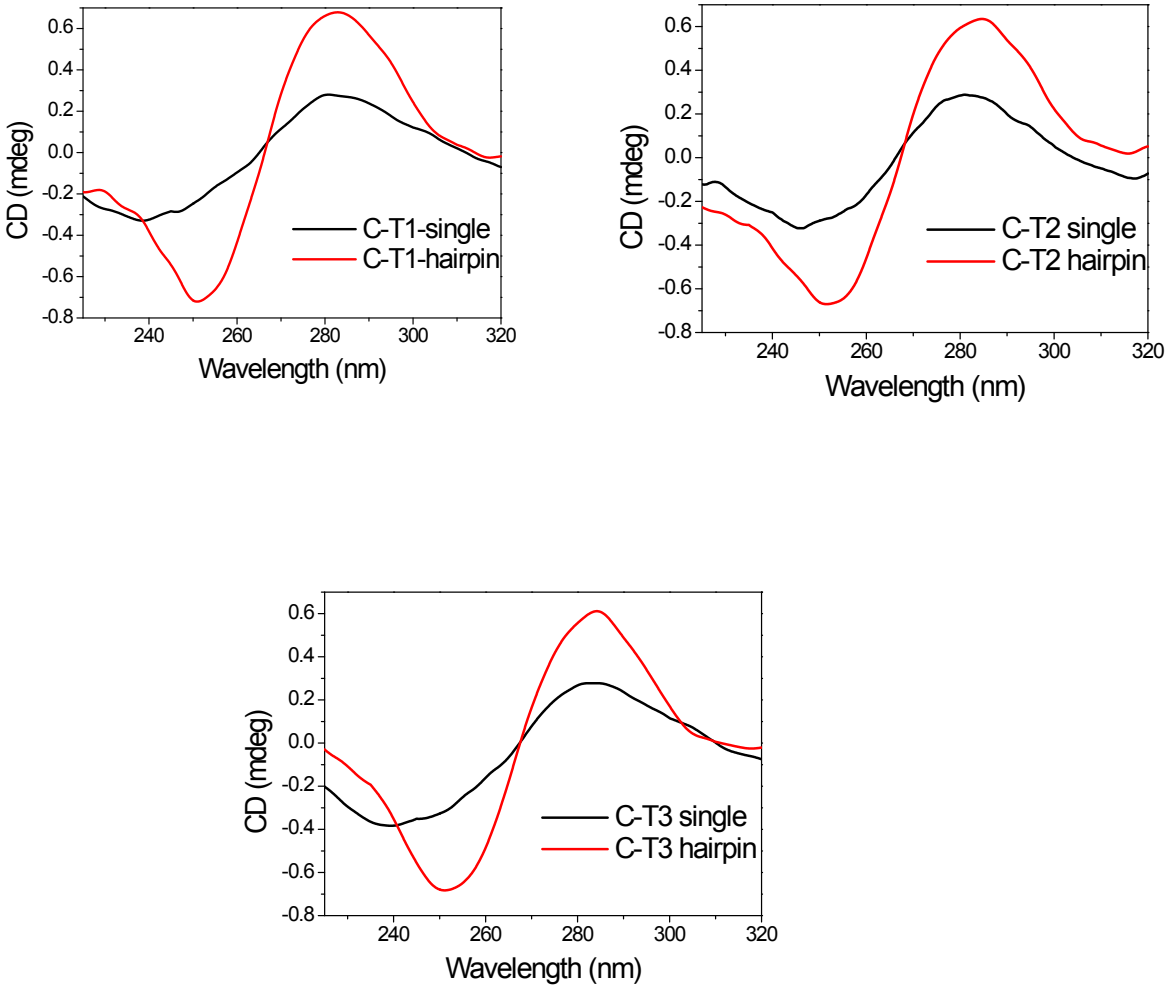


Figure S3. CD spectra of the full length products. All spectra were measured in the range 200–400 nm. All hairpin samples were prepared in 25 mM Trizma buffer (50 mM NaCl, 10 mM $MgCl_2$, 25 mM Tris-HCl, pH 7.5). DNA single-strand samples were prepared in only Tris buffer.

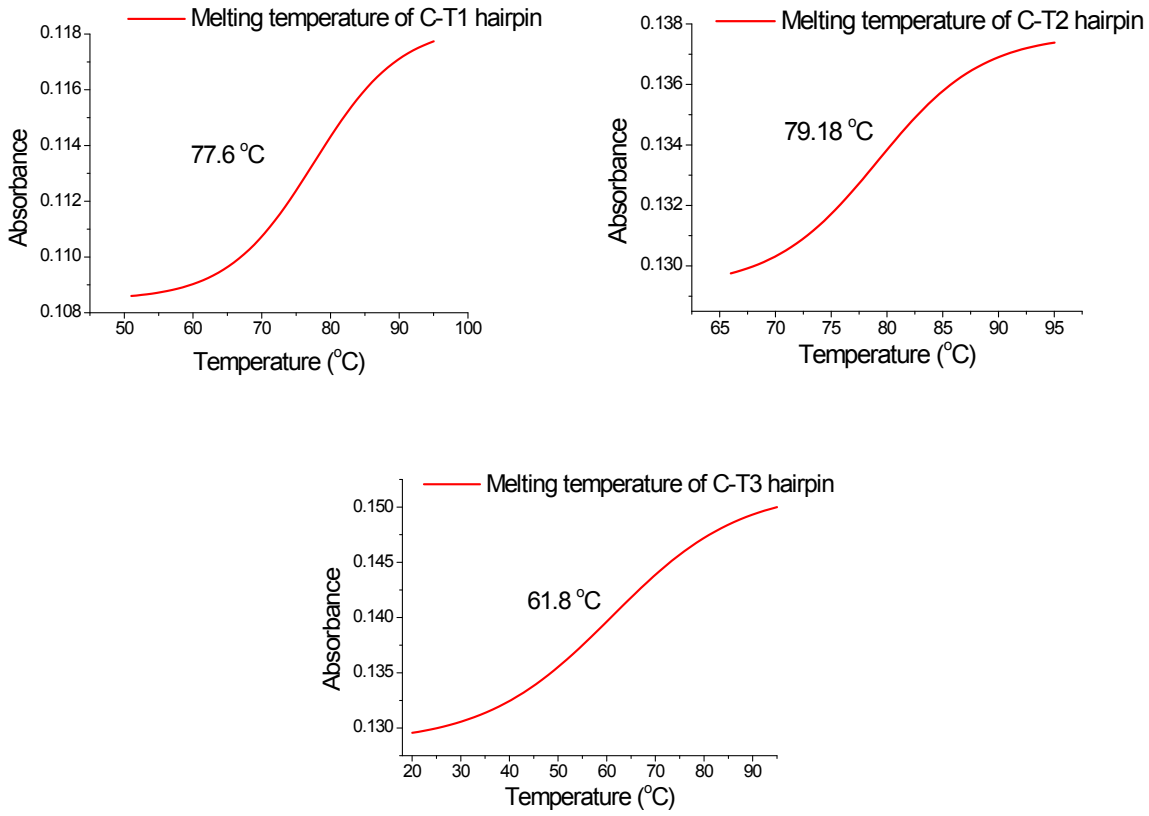


Figure S4. Melting temperatures of the full length products **C-T1**, **C-T2**, and **C-T3**. All samples were prepared in 25 mM Trizma buffer (50 mM NaCl, 19 mM $MgCl_2$, pH 7.2) at 25 °C.

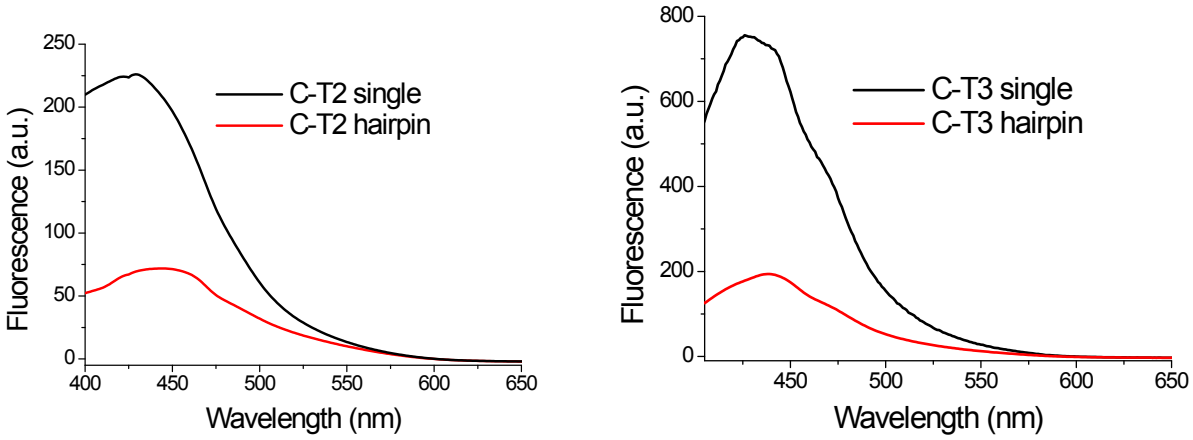
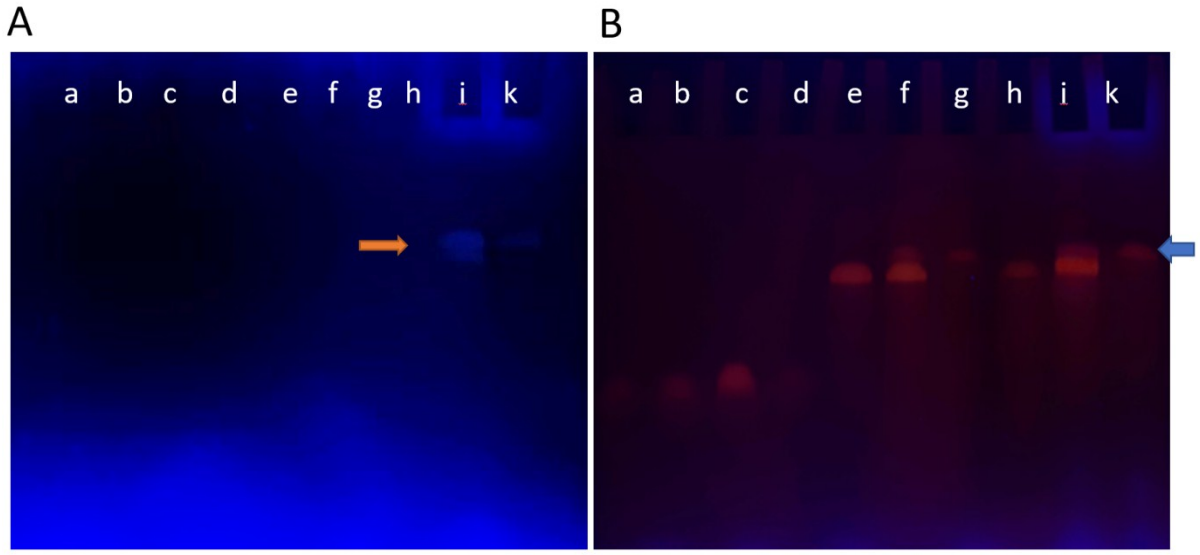
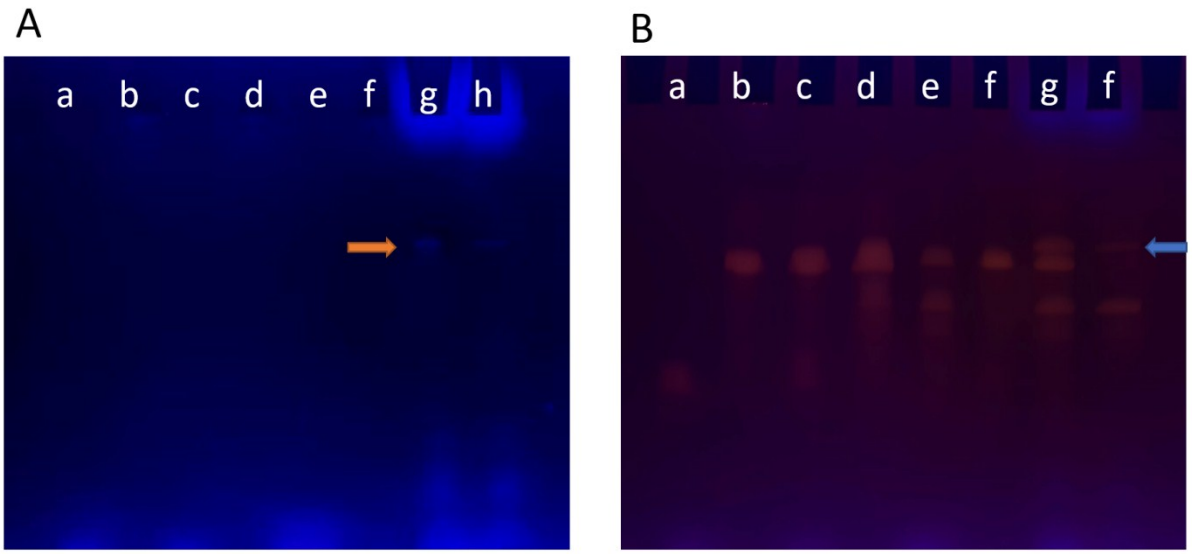


Figure S5. Fluorescence spectra of **C-T2** and **C-T3** in the single-strand and hairpin states. The hairpin DNA samples were prepared in 25 mM Trizma buffer (50 mM NaCl, 19 mM $MgCl_2$, pH 7.2) at 25 °C. The single-strand DNA samples were prepared in Tris buffer alone. The excitation wavelength was 386 nm.



Primer	5′-ATCCAAGGATGCATT-3′ (15 mer)
Template homoA-1	5′-[Phosphate]AAATGCATCCTTGGAT-3′ (16 mer)
Template homoA-2	5′-[Phosphate]AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATGCATCCTTGGAT-3′ (46 mer)
Template homoA-3	5′-[Phosphate]AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATGCATCCTTGGAT-3′ (47 mer)

Figure S6. Denaturing PAGE characterization of the proposed strategy using homoA-sequence. (**A**) Fluorescence and (**B**) staining (EtBr) images of the gel. (a) **Primer**; (b) **Template homoA-1**; (c) after incorporating one **dUazoTP** unit; (d) treating Lambda exonuclease after incorporating one **dUazoTP** unit; (e) **Template homoA-2**; (f) after extending with **dTTP**; (g) treating Lamda exonuclease after extending with **dTTP**; (h) **Template homoA-3**; (i) after incorporating one **dUpyrTP** unit; (k) full length product after treating with Lambda exonuclease. **Primer** (0.1 mM) and **Template homoA-1** (0.1 mM) were annealed to form a **Primer :Template homoA-1** duplex by heating to 95 °C and then slowly cooling to room temperature. nPfu DNA polymerase (2 U) was then added with 2 mM **dUazoTP** and then the system was incubated at 37 °C for 12 h; This sample was treated with Lambda exonuclease (5 U, Glycine/KOH buffer, 37 °C, 30 min) and then purified using a QIAquick nucleotide removal kit (**step 1**). The single strand product from step 1 was then annealed with **Template homoA-2** (0.1mM) by heating to 95°C and slowly cooling to room temperature. nPfu DNA polymerase (2 U) and 2 mM **dTTP** were added to this sample, which was then incubated at 37 °C for 12 h. The sample was also treated with Lambda exonuclease (5 U, Glycine/KOH buffer, 37 °C, 30 min) to cleave **template homoA-2** and purified using a QIAquick nucleotide removal kit (**step 2**). The single strand product from step 2 was then annealed with **Template homoA-3** (0.1mM) by heating to 95°C and slowly cooling to room temperature. nPfu DNA polymerase (2 U) and 2 mM **dUpyrTP** were added to this sample, which was then incubated at 37 °C for 12 h. Finally, the sample was treated with Lambda exonuclease (5 U, Glycine/KOH buffer, 37 °C, 30 min) to get the full-length product (**step 3**). All polymerase extensions were stopped by adding twice the amount of stop buffer (10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol, and 0.1% bromophenol blue in formamide). All product solutions were loaded onto a 20% denaturing polyacrylamide gel.



Primer	5′-ATCCAAGGATGCATT-3′ (15 mer)
Template G4-1	5′-[Phosphate]CCCTAACCCTAACCCTAACCCAAATGCATCCTTGGAT-3′ (37 mer)
Template G4-2	5′-[Phosphate]ACCCTAACCCTAACCCTAACCCAAATGCATCCTTGGAT-3′ (38 mer)

Figure S7. Denaturing PAGE characterization of the proposed strategy using C-rich sequence. (A) Fluorescence and (B) staining (EtBr) images of the gel. (a) **Primer**; (b) **Template G4-1**; (c) after incorporating one **dUazoTP** unit; (d) after extending with **dATP, dGTP**, and **dTTP**; (e) treating Lamda exonuclease after extending with **dATP, dGTP** and **dTTP**; (f) **Template G4-2**; (g) full length product with incorporation of **dUpyrTP**; (h) full length product after treating with Lambda exonuclease. **Primer** (0.1 mM) and **Template G4-1** (0.1 mM) were annealed to form a **Primer:Template G4-1** duplex by heating to 95 °C and then slowly cooling to room temperature. nPfu DNA polymerase (2 U) was then added with 2 mM **dUazoTP** and then the system was incubated at 37 °C for 12 h; this sample was purified using a QIAquick nucleotide removal kit (**step 1**). nPfu DNA polymerase (2 U) and 2 mM **dATP**, **dGTP**, and **dTTP** were added to this sample, which was then incubated at 37 °C for 12 h and purified using a QIAquick nucleotide removal kit. The sample was treated with Lambda exonuclease (5 U, Glycine/KOH buffer, 37 °C, 30 min) to cleave **template G4-1** (**step 2**). The single strand product from step 2 was then annealed with **Template G4-2** (0.1mM) by heating to 95°C and slowly cooling to room temperature. nPfu DNA polymerase (2 U) was then added with 2 mM **dUpyrTP** and then the system was incubated at 37 °C for 12 h. After that, product was purified by using a QIAquick nucleotide removal kit. Finally, the sample was treated with Lambda exonuclease (5 U, Glycine/KOH buffer, 37 °C, 30 min) to get the full-length product (**step 3**). All polymerase extensions were stopped by adding twice the amount of stop buffer (10 mM EDTA, 10 mM NaOH, 0.1% xylene cyanol, and 0.1% bromophenol blue in formamide). All product solutions were loaded onto a 20% denaturing polyacrylamide gel.

Full-length product sequence	Yield (%)
C-T1	36.78
C-T2	25.00
C-T3	16.17
Product using Template homoA-1, A-2, and A-3	23.97
Product using Template G4-1 and Template G4-2	14.41

Table S2. The yields of the labeled full-length products were calculated using the fluorescence intensity of bands in PAGE after staining with Ethidium bromide (**Fig. 1, Fig S2, Fig S6, Fig S7**). Comparison fluorescence intensities between full-length product bands and template bands were measured by using Gel Analyzer software (http://www.gelanalyzer.com/) .

References:

1. C. F. Barbas, D. R. Burton, J. K. Scott and G. J. Silverman, *Cold Spring Harbor Protocols*, 2007, **2007**, pdb. ip47.
2. S. R. Gallagher, *Current protocols in neuroscience*, 2011, **56**, A. 1K. 1-A. 1K. 14.
3. J. C. Sutherland, D. C. Monteleone, J. Trunk and G. Ciarrocchi, *Analytical biochemistry*, 1984, **139**, 390-399.